AD-783 073

A GENERALIZED DISTANCE' ESTIMATION PROCEDURE FOR INTRA-URBAN INTERACTION

A. Charnes, et al

Texas University

Prepared for:

Office of Naval Research
National Science Foundation

April 1974

DISTRIBUTED BY:

NTIS
National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

Unclassified

Security Classification

AD 783073

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Center for Cybernetic Studies The University of Texas	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

A 'Generalized Distance' Estimation Procedure for Intra-Urban Interaction

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

A. Charnes
Kingsley E. Haynes
Fred Y. Phillips

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
April 1974	18	14

8a. CONTRACT OR GRANT NO. N.S.F.(Rann) Grant No. GI-34870X; ONR # N00014-67-A-0126-0009 b. PROJECT NO. NR 047-021 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Center for Cybernetic Studies Research Report CS 171 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Office of Naval Research (Code 434) Washington, D.C.

13. ABSTRACT

The estimation of urban and regional travel patterns has been a necessary part of current efforts to establish land use guidelines for the Texas coastal zone [7]. Emerging from this research are several theoretical [4,12] and computational advances relating to the gravity model of spatial interaction. This paper details our computational experience with travel estimation within Corpus Christi, Texas, using a new convex programming approach of Charnes, Raike and Bettinger [4]. It is found that available estimation techniques necessarily result in non-integer solutions. A mathematical device is therefore introduced which assures integer-values estimates. Implications and extensions of the work are indicated.

Unclassified

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Gravity Model						
Integer Programming						
Statistical Estimation						
Transportation						
Convex Programming						
Spatial Interaction						
Regional Management						

DD FORM 1 NOV 65 1473 (BACK)

S/N 0102-014-6800

ia

Unclassified

Security Classification

Research Report
CS 171

A 'GENERALIZED DISTANCE' ESTIMATION PROCEDURE FOR INTRA-URBAN INTERACTION

by

A. Charnes[1]
Kingsley E. Haynes[2]
Fred Y. Phillips[3]

April 1974

[1]Dr. Charnes is University System Professor of The University of Texas System.

[2]Dr. Haynes is Associate Professor at The Lyndon B. Johnson School of Public Affairs, and a member of the Department of Geography.

[3]Mr. Phillips is a Ph.D. candidate in Management Science.

This research was partly supported by N.S.F. (RANN) Grant No. GI-34870X "Establishment of Environmental Guidelines for Environmental Management" and by Project No. NR 047-021, ONR Contract N00014-67-A-0126-0009 with the Center for Cybernetic Studies, The University of Texas.

CENTER FOR CYBERNETIC STUDIES

A. Charnes, Director
Business-Economics Building, 512
The University of Texas
Austin, Texas 78712

INTRODUCTION

As reported elsewhere [3, 6, 11] a large scale modelling effort is underway for the purpose of developing land use and environmental guidelines in the Texas Coastal Zone. The land use portion of this project requires the spatial allocation of commercial, retail and residential elements of the Corpus Christi region under a series of policy constraints. As an input to this analysis it has been important to determine the character of spatial interaction for a series of activities not the least of which was the intra-urban traffic pattern. A set of experiments have been conducted on various interaction estimation procedures and the results of one of these experiments is reported below. The importance of this experiment lies in the demonstrated applicability of an extremal method of statistical estimation based on the statistical information-theoretic approach outlined by Charnes, Raike and Bettinger (4). The demonstration rests on a new mathematical device, herein introduced for the first time. This device incorporates our desire for integer-valued estimates of the traffic flows, and in fact guarantees integer solutions. Past literature and computation has failed to come to grips explicitly with this problem. We believe this to be the first full scale application of the approach and feel that some of the issues faced may be of value to other spatial analysts concerned with interaction estimation problems.

I. THE ESTIMATION PROBLEM

The Texas Highway Department [13] has compiled and up-dated origin-destination (O-D) data for the fifty-six traffic districts within Corpus Christi, both for vehicle trips and for passenger trips, with each trip type disaggregated by vehicle type. Summing these data yielded the aggregate trip type "TOTAL INTERNAL PERSON TRIPS-ALL MODES OF TRAVEL." We considered our first estimation task to be the district-to-district trip distribution of this trip type.

We wished to estimate a set $\{t_{ij}; i, j=1, \ldots, 56\}$, where t_{ij} is the number of persons traveling from district i to district j in a day's time. The t_{ij} were to satisfy a "gravity equation" of the form

$$t_{ij} = \frac{\tau_i \sigma_j}{g(K_{ij})} \tag{1}$$

where K_{ij} is the "generalized distance" between districts i and j, τ_i and σ_j (i, j=1, ..., 56) are "gravity potentials", and g is a function whose form is determined by the estimation procedure. The O-D data and the K_{ij} are known constants, and are inputs to the estimation procedures. In the present effort, the K_{ij}'s are a function of the road distance between districts and of the speed limits on these roads; they were approximated by inspection of maps of the city.

The trip distribution estimates from this procedure will be among the inputs to the retail location sector of the Project [7].

II. SOLUTION METHODS CONSIDERED

D'Esopo and Lefkowitz [5] describe an iterative procedure leading to estimates of the τ_i and σ_j of equation (1). This algorithm is not guaranteed to always converge, and hence was dismissed as unreliable.

Wilson's [6] method uses one of the best-known of the "entropy-maximizing" approaches to travel pattern estimation. The two principal objections to this approach were (a) its extravagant data requirements, which called for information beyond that available from the census, and (b) the unusual and nontraditional form of the function g of equation (1) that results from the algorithm.

Charnes, Raike, and Bettinger showed (a) that the t_{ij} satisfying (1) may be obtained by an extremal principle of convex programming form, and (b) that this principle is equivalent to a standard statistical estimation procedure of the information-theoretic variety [10] -- a hitherto unknown fact. The CRB procedure has none of the limitations of Wilson's or the D'Esopo-Lefkowitz methods, and is computationally practicable.[1]

III. THE SOLUTION PROCEDURE

Using the Kuhn-Tucker optimality conditions Charnes, Raike, and Bettinger [4] prove the following Theorem. Consider the convex programming problem; minimize $\sum_{i \neq j} (t_{ij} \ln K_{ij} t_{ij} - t_{ij})$ (2)

---

[1]For a more detailed comparison of the latter two solution methods see [8] and [12].

subject to

$$\sum_{\substack{i=1\\ i\neq j}}^{n} t_{ij} = D_j \qquad i=1, \ldots, n$$
$$\sum_{\substack{j=1\\ i\neq j}}^{n} t_{ij} = O_i \qquad j=1, \ldots, n \tag{3}$$

and all $t_{ij} \geq 0$.

If a travel distribution of positive t_{ij} satisfies the gravity equation (1) for some set of potentials τ_i and σ_j and for some distance function K_{ij} and if these t_{ij} satisfy the origin-destination requirements (3), then these t_{ij} solve the convex programming problem (2), (3) optimally. Conversely, if a distribution $\{t^*_{ij}\}$ solves (2), (3) optimally, then potentials τ_i, σ_j exist such that equation (1) is satisfied with $g(K_{ij}) = K_{ij}$.

This optimization is shown to yield the trip distribution which is statistically "least distinguishable", in terms of the statistical information-theoretic measure [4, 10], from the ease-of-travel distribution.[2] An additional property of an exact optimal solution to such a convex programming problem is that a large fraction of the optimal t_{ij} <u>must</u> be non-integer. This fact has gone unmentioned in all past literature.

Since the function (2) is convex and separable in the t_{ij}, it may be approximated by a piecewise linear function of the t_{ij} [1, p. 348]. Thus transformed, the problem would take the form of a "capacitated distribution problem." This special structure can be solved with suitable special

---

[2]Ease-of-travel is defined as the inverse of the generalized travel cost, normalized to a 0-1 interval.

computer codes two orders of magnitude faster than with general purpose linear programming codes. Facing squarely the fact that we require integer values for the estimated t_{ij}, we have developed a new two-piece approximation which guarantees (for extreme point solutions) that the t_{ij} will be integers.

One difficulty in the capacitated distribution approach lay in the fact that most computer codes suitable for such problems will accept only integer data. Due to the logarithmic terms in the functional (2), truncating the coefficients at the decimal point would have resulted in catastrophic rounding errors. Thus it was necessary to scale all of the input data upward by a factor of 10^5 and, of course, scale the resulting optimal t_{ij} downward by the same factor.

Initially an "out-of-kilter" algorithm code was employed for solving the distribution problem. A minor complication at this stage was that the input and output formats of the out-of-kilter code were not equipped to handle integers of the magnitude of the "scaled-up" input data. Hence these formats had to be modified.

The all-in-core program was run in 422 seconds, and required all of the core memory space available for user programs on the CDC 6600 at the University of Texas. The explanation for these unusually large time and core requirements may lie in the fact that (a) all districts are available as destinations for each district of origin, and (b) the 2-piece

linear functional necessitates two links between each node (district) of the network. Thus the network is extremely connected, as well as dense, compared to that of the typical capacitated distribution problem.

Because of the extraordinary time required by the out-of-kilter algorithm, the problem was run on a "primal" algorithm code of Dr. Darwin Klingman's.[3] An optimal solution was reached in only nine seconds with this code. Moreover, the core requirement was reduced by three-fourths.[4]

IV. DERIVATION OF THE DIOPHANTINE PIECEWISE LINEAR APPROXIMATION

Recalling the nonlinear equation system (2), (3), we transform the variables t_{ij} as follows:

$$t_{ij} \to \delta_{ij}^{1} + \delta_{ij}^{2} = \sum_{k=1}^{2} \delta_{ij}^{k} \tag{4}$$

The origin-destination equations (3) become

$$\sum_{\substack{i=1 \\ i \neq j}}^{n} \sum_{k=1}^{2} \delta_{ij}^{k} = D_j \qquad j=1, \ldots, n \qquad \text{(3a)}$$

and

$$\sum_{\substack{j=1 \\ i \neq j}}^{n} \sum_{k=1}^{2} \delta_{ij}^{k} = O_i \qquad i=1, \ldots, n \qquad \text{(3a)}$$

with all $\delta_{ij}^{k} \geq 0$.

---

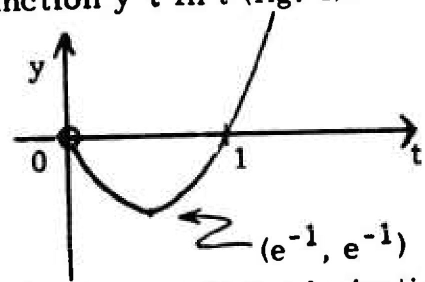
[3] We wish to thank Richard Barr of the Center for Cybernetic Studies for making the out-of-kilter code available and for patiently performing the necessary alterations, and Dr. Darwin Klingman for executing the primal code calculation.

[4] Problems much larger than the current one may be solved by any of the many available distribution codes which are not executed wholly in-core.

The functional (2) to be minimized is

$$\sum_{i\neq j}(t_{ij} \ln k_{ij}\, t_{ij} - t_{ij}) = \sum_{i\neq j}(t_{ij} \ln k_{ij} + t_{ij} \ln t_{ij} - t_{ij}) \tag{5}$$

The only nonlinear term on the right-hand side of (5) is of form $(t \ln t)$. Graphing the function $y = t \ln t$ (fig. 1):



(The function has a singularity at $t=0$; the derivative is infinite.)

Figure 1

It is this function that must be approximated by a piecewise linear function. Piecewise linearity is achieved by introducing upper bounds on the δ_{ij}^{1}. The optimal (extreme point) t_{ij} will then be integral or non-integral respectively as the bounds are integral or non-integral. The procedure naturally suggested by the shape of the curve, i.e., to place the "knee" of the approximation at (e^{-1}, e^{-1}) and to bound the δ_{ij}^{1} at e^{-1}, is therefore useless for the problem at hand, because the resulting optimal t_{ij} will not be integers.

Bearing in mind that only integer results are competitive, we can modify the functional over the domain of its non-integral values. We introduce the modification shown in Figure 2. It guarantees integrality and an appropriate functional behavior over integer values.

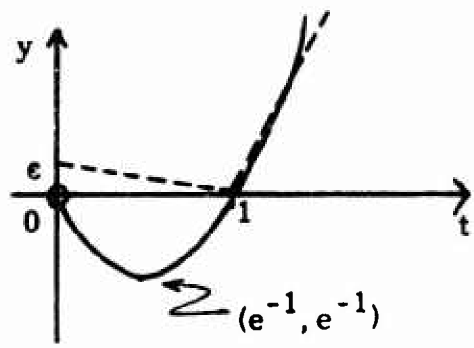


Figure 2

Here, the slope of the linear function is $-\epsilon$ for $0 \le t < 1$, and R for $1 < t$, where $0 < \epsilon << 1$ and $R = \min \{\max_i O_i, \max_j D_j\}$.

Thus, if we observe the condition $0 \le \delta^1 \le 1$, then $-\epsilon\delta^1 + R\delta^2 \approx t \ln t$, where $t = \delta^1 + \delta^2$. Therefore the functional (5) equals, under the transformation (4),

$$\sum_{i \neq j} [(\delta_{ij}^1 + \delta_{ij}^2) \ln K_{ij} + (\delta_{ij}^1 + \delta_{ij}^2) \ln (\delta_{ij}^1 + \delta_{ij}^2) - (\delta_{ij}^1 + \delta_{ij}^2)]$$

$$\approx \sum_{i \neq j} [(\delta_{ij}^1 + \delta_{ij}^2) \ln K_{ij} + (-\epsilon\delta_{ij}^1 + R\delta_{ij}^2) - (\delta_{ij}^1 + \delta_{ij}^2)], \qquad (6)$$

which is clearly linear in the δ_{ij}^k. It remains to minimize the right-hand side of equation (6) subject to the conditions (3a) and the further condition $0 \le \delta_{ij}^1 \le 1$ for all i, j.

V. RESULTS AND INTERPRETATIONS

It will be recalled that the first run was for the trip type "total internal person trips, all modes of travel" (Figure 3). We anticipate further applications of the method developed, including (a) other trip types of less aggregated character, and (b) inclusion of the major traffic

Corpus Christi Travel Distribution Estimate: 56 Internal Districts

Total Internal Person-trips; All Modes of Travel

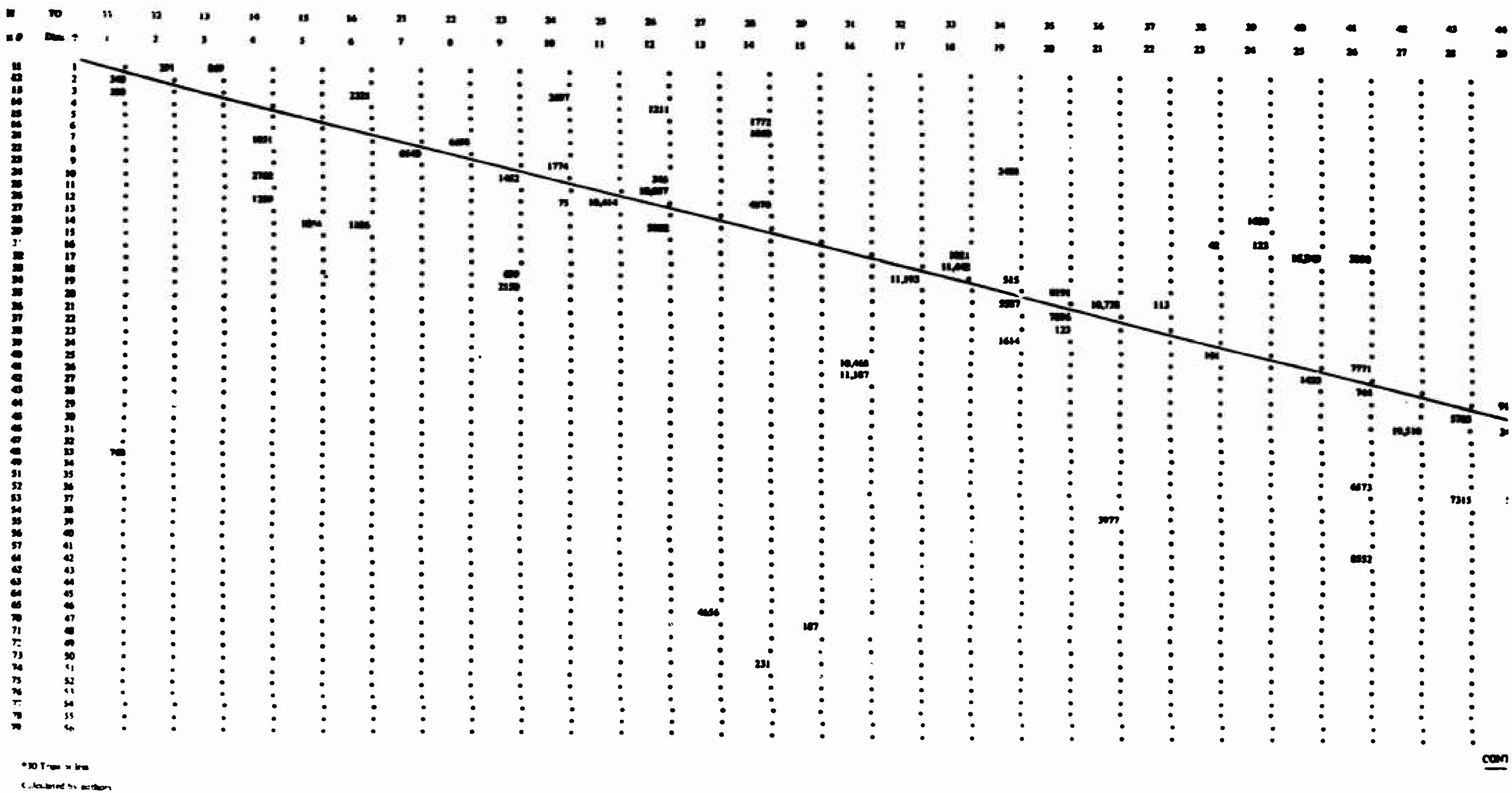


FIGURE 3

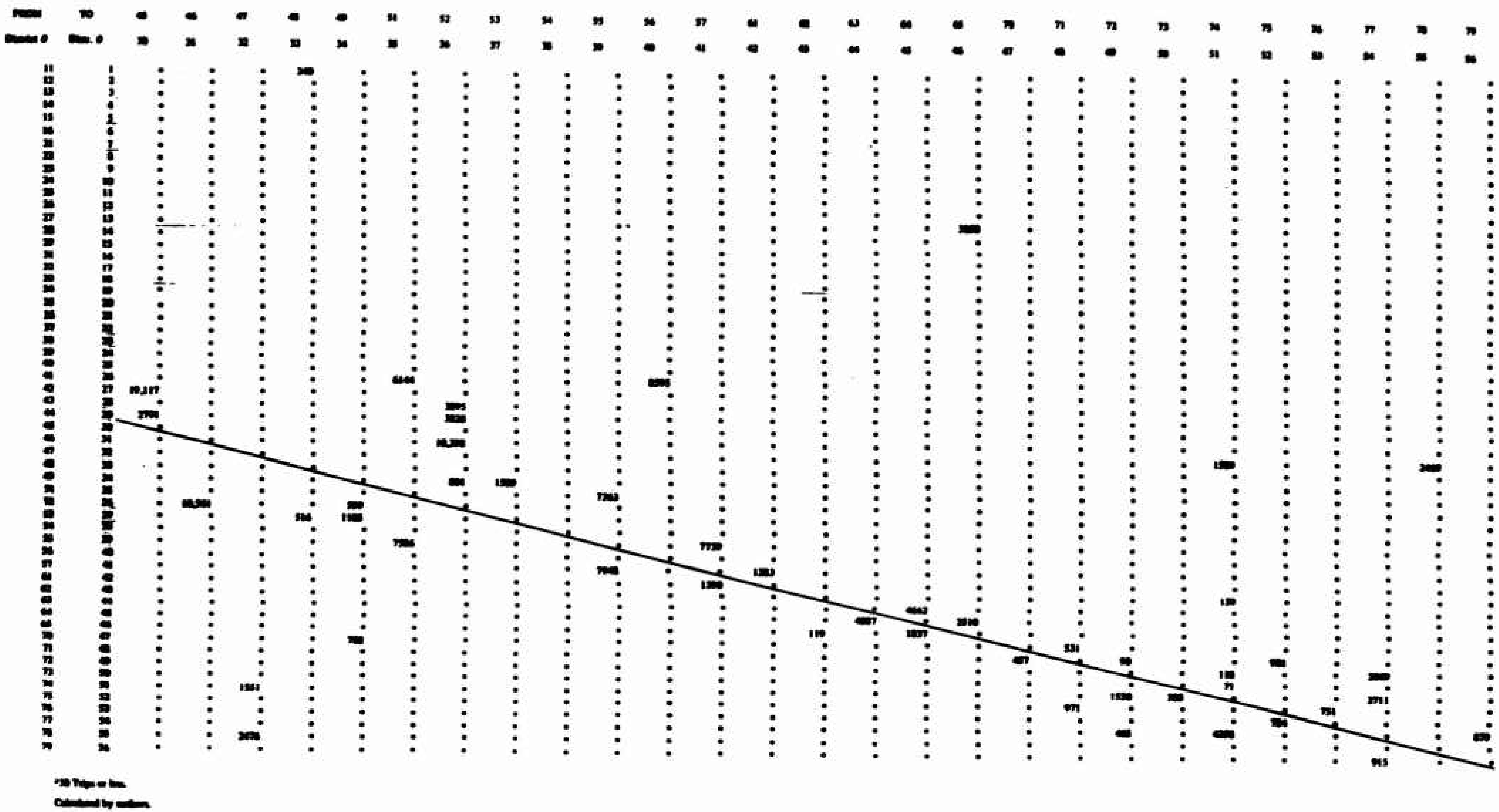


FIGURE 3

arteries ("external stations") entering and leaving the continuously urbanized area.

The travel distribution estimate obtained from the procedure outlined above had some dominating characteristics. Traffic volume was concentrated on relatively few interdistrict links. Other links were assigned token transfers many orders of magnitude less than those of the major links. Further, a pair of adjacent districts are often seen to be connected by a major link, exchanging approximately equal numbers of travelers.

A potentially important result of this study is the observation that the distance between districts (as we have measured it) apparently exerted strong influence on the distribution estimate, despite the fact that in equation (1), $g(K_{ij}) = K_{ij}$. Recall that in many previous studies, an effort is made to represent $g(K_{ij})$ as K_{ij}^{α}, where $\alpha > 1$, giving distance a heavier weight in the estimate. This may prove to have been a superfluous endeavor.

VI. CONCLUSION

Further extension of the work of Charnes, Raike, and Bettinger [4] is in progress, including a new direct characterization of the potentials τ_i, σ_j through optimal solution of an unconstrained convex programming problem. Corresponding development of a new algorithm for the t_{ij}

follows. These characterizations stem from the dual of a variant of an extended geometric programming problem put forward in another context by Charnes and Cooper [2].

REFERENCES

1) A. Charnes and W. W. Cooper, Management Models and Industrial Applications of Linear Programming. John Wiley, New York (1961).

2) A. Charnes and W. W. Cooper, "An Extremal Principle for Accounting Balance of a Resource Value-Transfer Economy: Existence, Uniqueness, and Computation," Rendiconti di Academia Nazionale dei Lincei (April 1974).

3) A. Charnes, K. E. Haynes, J. E. Hazleton and M. J. Ryan, "The Texas Coastal Zone Project" NATO Conference, Mathematical Analysis of Decision Problems in Ecology, Istanbul, Turkey, (July, 1973).

4) A. Charnes, W. Raike, and C. O Bettinger, "An Extremal and Information-Theoretic Characterization of Some Interzonal Transfer Models." Socio-Economic Planning Sciences, Vol. 6, pp. 531-537 (1972).

5) D. A. D'Esopo and B. Lefkowitz, "An Algorithm for Computing Interzonal Transfers Using the Gravity Model." Operations Research Vol. 11, pp. 901-906 (1963).

6) E. Gus Fruh and Joe C. Moseley II, "Management and Texas Coastal Resources" National Science Foundation (RANN) Symposium (November, 1973) pp. 142-146.

7) K. E. Haynes and J. E. Hazleton, Establishment of Operational Guidelines for Texas Coastal Zone Management: Interim Report on Economics and Land Use. The University of Texas at Austin for N.S.F. (RANN). (May, 1973).

8) K. E. Haynes and F. Phillips, "Information Theoretic Approaches to Spatial Interaction" (Mimeograph, October 30, 1972) University of Texas, Department of Geography.

9) D. Huff, "Defining and Estimating a Trading Area," J. Marketing, pages 34-38 (1964).

10) S. Kullback, Information Theory and Statistics. John Wiley & Sons, New York (1959).

11) "Managing Coastal Lands" Mosaic, Vol. IV, (Summer, 1973), pp. 27-32.

12) F. Phillips and Gerald M. White, "Extremal Approaches to Estimating Spatial Interaction," Center for Cybernetic Studies Research Report CCS 168, University of Texas at Austin, April 1974.

13) Texas Highway Department, Planning Survey Division, Corpus Christi Metropolitan Transportation Study (Origin Destination Survey) in cooperation with U. S. Department of Commerce, Bureau of Public Roads and the City of Corpus Christi, 1961. (Revised and up-dated 1964, 1968, and 1972).

14) A. G. Wilson, "A Statistical Theory of Spatial Distribution Models." Transportation Res. Vol. 1, pp. 253-269, (1967).